
CST1204: Introduction to Databases

Jun Shan
Adjunct Professor
City Tech, CUNY

Week 12 Session 2

11/13/2019

Homework Review

- Chapter 5 exercise questions for Colonial Adventures 1-5

Mixed Usage of WHERE Clause

- Mixed usage of WHERE clause: Readability issue

```
SELECT r.REP_NAME  
FROM CUSTOMER c, REP r  
WHERE c.REP_NUM = r.REP_NUM  
AND c.CUST_NUM IN (1, 2, 3)
```

JOIN ON

- Separate join from filtering conditions
- Easier to see join linkage when joining multiple tables
- ANSI 92 standard (joining by WHERE is ANSI 89 standard).

```
SELECT r.REP_NAME  
FROM CUSTOMER c,  
      REP r  
WHERE c.REP_NUM = r.REP_NUM  
AND c.CUST_NUM IN (1, 2, 3)
```



```
SELECT REP_NAME  
FROM CUSTOMER c  
JOIN REP r  
      ON c.REP_NUM = r.REP_NUM  
WHERE c.CUST_NUM IN (1, 2, 3)
```

JOIN ON

- ANSI notation of join

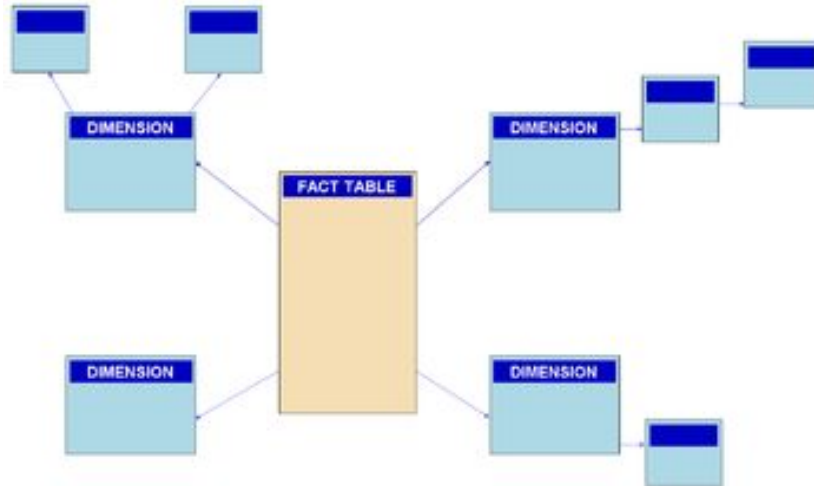
```
SELECT <qualified column list>
FROM <table 1> <alias 1>
JOIN <table 2> <alias 2>
    ON <matching column 1> = <matching column 1>
WHERE <boolean expression>
GROUP BY <GROUP BY column list>
HAVING <boolean expression>
ORDER BY <ORDER BY column list>
```

Multi Table Join

- Star join: Join from one table to multiple tables
- Chain join: Join from table A to table B, then from table B to table C ...
- Snowflake join: A combination of star join and chain join
- Avoid looping joins

Star/Snowflake Schema

- Most common usage of SQL query: star/snowflake schema



JOIN vs Normalization

- There are two types of data usage:
 - Operational (transactional):
 - Fast. Single entry. Requires integrity.
 - Requires normalization to avoid complexity and redundancy.
 - Analytical:
 - Historical statistics. Involving a variety of information
 - Needs to join data from multiple places to perform complex transformation/calculation

Join with GROUP BY

- Execution order:
 - FROM ... JOIN ... WHERE will create a detailed dataset
 - GROUP BY on top of this detailed data set will form an aggregated dataset over the joined tables
 - Use HAVING to filter on aggregated dataset

How to Write a SELECT Statement

1. Starting from the detailed data set
 - Starting from FROM: The table with most join keys should be in FROM clause (usually it is also the table with lowest level of details)
 - From FROM to JOIN: Add table name in the JOIN clause, then add join key in the ON clause.
 - Each new table is a independent JOIN. Join one table at a time.
 - Add WHERE condition

How to Write a SELECT Statement

2. Apply aggregation

- Add GROUP BY for each “for each”
- Add the same columns to SELECT
- Add required aggregation functions to SELECT
- Add HAVING if needed

3. Apply ORDER BY

Hands On

- Chapter 5 exercise questions for TAL Distributors 7.
 - How many order lines for each item?
 - Average number of items of each line
- Chapter 5 exercise questions for Colonial Adventure 3-5

JOIN execution

- Sort both tables by join key. Match and merge.

CUSTOMER	
Customer_Name	Rep_Num
Jeff	101
Maria	101
Rachel	101
Jose	102
Fred	102



REP	
Rep_Num	Rep_Name
101	Eric
102	Bill



Join result		
Customer_Name	Rep_Num	Rep_Name
Jeff	101	Eric
Maria	101	Eric
Rachel	101	Eric
Jose	102	Bill
Fred	102	Bill

Filtering Data using Another Table

- Method 1: Use WHERE IN (subquery) clause
- Method 2: Join two tables together. Filter on the joined dataset
- Method 3: Use WHERE EXISTS ()
 - Subquery with correlation condition between main query and subquery (which is the ON clause in JOIN between two main query tables)

EXISTS

- Check the result of a correlated subquery
- Process
 - Outside query is executed row by row
 - For each row, go through the inner query row by row, matching outside query's current row

JOIN vs IN vs EXISTS

- Similarity in execution: Matching values in another table
 - SELECT contains information from both tables: JOIN
 - SELECT only from one main table and condition is not related to main table: IN
 - SELECT only from one main table and condition is related to main table: EXISTS
 - In most cases, if you can use IN or EXISTS, you can also use JOIN

Hands On

- Chapter 5 exercise questions for TAL Distributors 4-5 using
 - JOIN ... ON ...
 - IN
 - EXISTS

Homework

- Chapter 5 exercise questions for Colonial Adventure Tours 6-9
- Chapter 5 exercise questions for Solmaris Condominium Group 4-8